

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4162, and to the Office of Management and Budget, Paperwork Reduction Project 0704-0188, Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank) 2. REPORT DATE 22/2/1997 3. REPORT TYPE AND DATES COVERED FINAL TECH RPT, 01 Apr 93-30Sep 96

4. TITLE AND SUBTITLE Theory and Algorithms for Memory Based Inference 5. FUNDING NUMBERS F49620-93-1-0266

6. AUTHOR(S) George Cybenko

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Thayer School of Engineering
Dartmouth College
Hanover, NH 03755

0325

8. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) 10. SPONSORING/MONITORING AGENCY REPORT NUMBER

Air Force Office of Scientific Research/NM
Bolling AFB, DC 20332

F49620-93-1-0266

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT 12b. DISTRIBUTION CODE

Approved for public release; distribution
unlimited.

13. ABSTRACT (Maximum 200 words)

The objectives of this project were to develop new methods for analyzing memory-based techniques in machine learning and general inference problems. The project successfully developed new mathematical results explaining when memory-based methods are useful. Several applications were developed and studied in this project. A system for building transportable agents was developed and several information processing systems have been based on it. Technology has been successfully transitioned to both DOD and commercial entities.

14. SUBJECT TERMS 15. NUMBER OF PAGES

Machine learning, intelligent agents, estimation.

5

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT 18. SECURITY CLASSIFICATION OF THIS PAGE 19. SECURITY CLASSIFICATION OF ABSTRACT 20. LIMITATION OF ABSTRACT

Unclassified

Unclassified

Unclassified

UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
GPO:1992

53-85

DTIC QUALITY INSPECTED 2

TOTAL P.02

19980421 147

Final Project Report: Theory and Algorithms for Memory Based Inference

Grant Title: Theory and Algorithms for Memory Based Inference

AFOSR Grant Number: #F49620-93-1-0266

Project Duration: 1 April 93 - 1 September 96

PI Name: George Cybenko

PI Address: Thayer School of Engineering, Dartmouth College, Hanover, NH 03755

PI Phone: (603) 646-3843

PI Fax: (603) 646-2277

PI email address: gvc@dartmouth.edu

Objectives

The research objectives of this project have been to study the theoretical and algorithmic underpinnings of memory-based learning and inference and then to explore uses of the methodology in timely and important applications. The use of memory-based methods has been growing in many application areas because of two reasons: 1. large fast computer storage has become more affordable and; 2. parametric modeling techniques are often difficult to train and use on problems with large data sets. The initial focus of the effort was to discover the mathematical foundations for memory-based methods and to develop algorithms that would allow efficient use of large memories for pattern analysis and recognition. The application area that received the most attention in this project has been networked information processing using transportable agent systems.

Accomplishments

Research supported by this grant has resulted in advances in the theoretical aspects of memory-based learning. One of the original goals was to identify pattern analysis problems which could be solved by memory-based as opposed to parametric methods. Existing work on high-dimensional memory based methods has been relatively pessimistic about their utility in spite of experiments which showed that they worked well on real problems. In our work, we have shown that the concentration of data plays an important role in high dimensional real-world pattern recognition applications and measures of concentration provide metrics for how well memory-based methods work.

In the application of these basic results to distributed information management, the project has developed a large number of software systems for supporting text retrieval and transportable information agents. The agent system is based on the Tool Control Language, now an open system actively supported by SUN Microsystems. Called Agent Tcl, it has been downloaded by several hundred users at this time.

Agent Tcl is a transportable agent system with comprehensive functionality. It has been used in a variety

of applications including CAD mechanical parts searching, medical records retrieval, and tactical picture information processing. This last application has been developed by Lockheed Martin for the US Army. Their system has been demonstrated at Fort Bragg and other Army facilities.

The Lockheed Martin effort is using Agent Tcl to develop information agents that can assist Army decision-makers in low bandwidth battlefield situations. Their information agents move within a computer network that is vulnerable to changes in its nodes and connectivity retrieving data from remote sources, processing it and presenting summaries to battlefield commanders. Other serious defense and commercial users of the Agent Tcl system include:

- Mitre Corporation
- Network Operations Technologies Lab, NYNEX Science and Technology
- Siemens Corporate Research, Germany
- Telecom Bretagne, Network & Multimedia Dep. (RSM)
- AT&T Research.

The project has made several advances in memory-based methods and transportable information agent systems. Specific accomplishments have been:

- An analysis of conditions under which memory-based methods are effective for solving machine learning and associative memory processing problems.
<http://www.dartmouth.edu/~gvc/prague-ieee.ps>
- New results on on-line reinforcement learning using Q-learning to show that simultaneous learning and optimality are possible. <http://www.dartmouth.edu/~gvc/oxford.ps>
- A fully mobile agent programming system built on top of the Tool Control language (Tcl), called Agent Tcl. <http://www.cs.dartmouth.edu/~agent/papers/chapter.ps> is an overview and <http://www.cs.dartmouth.edu/~agent> is the comprehensive Web site.
- Analytic results relating agent performance with the architecture of the decision structure used. This appears to be the first attempt to model and study the relationship between automated distributed decision-making architectures and problem solution accuracy.
<http://www.dartmouth.edu/~gvc/uncertainty.ps>
- Demonstrations of the Agent Tcl technology to diverse application domains has been extremely successful, suggesting that the concept is robust and highly adaptable. See
<http://www.cs.dartmouth.edu/~agent/> and <http://comp-engg-www.dartmouth.edu/~3d/paper.ps>

Other Transitions

Basic research on this project has led to a World Wide Web application called the Informant (<http://informant.dartmouth.edu>) which has received numerous positive reviews in the popular literature as well as licensing requests from third parties. There are now over 13,000 registered users of this application world wide. Users include DOD personnel and agencies as well as other government users. A license arrangement with the National Air Intelligence Center (NAIC) is currently being discussed.

Several commercial software companies have approached the project regarding licensing and other commercialization plans for the Agent Tcl but they are in early stages of discussion.

Publications

1. On the unreasonable effectiveness of memory-based methods, in *Proceedings of IEEE Workshop on Large Scale Computation*, Prague, CZ, pp. 1-9, 1993. (Winner of best paper award)
2. Signal Processing Models for Information Retrieval, in *Proceedings of COST '94 Workshop on Emergent Techniques in Signal Processing and Communications*, VIGO, Spain, pp. 34-45, 1994.
3. Information Theoretic Principles of Agents, (with Robert Gray, Alexy Khrabrov and Yunxin Wu), in *CIKM Workshop on Intelligent Information Agents, Third International Conference on Information and Knowledge Management (CIKM '94)*, 1994, Gaithersburg, Maryland, Editors, Yannis Labrou and Tim Finin.
4. Information Agents as Organizers, (with Robert Gray, Alexy Khrabrov and Yunxin Wu), in *CIKM Workshop on Intelligent Information Agents, Third International Conference on Information and Knowledge Management (CIKM '94)*, 1994, Gaithersburg, Maryland, Editors, Yannis Labrou and Tim Finin.
5. Perfect reconstruction filtering with unitary operators and projections, (with M. Stewart), *Linear Algebra and its Applications*, Vol. 208, pp. 97-133, 1994.
6. Theory of Just-in-Time Modeling, in *From Estimation to Learning*, edited by S. Bittanti, Springer Verlag, 1996.
7. The future of information technology, in Optoelectronics for Data Communications, (edited by U. Osterberg and R. Lansky), Academic Press, pp. 311-320, 1995.
8. Signal Processing in The National Information Infrastructure, (with R. Gray et al.), National Science Foundation Report, 1995.
9. Q-Learning: A tutorial with extensions, to appear in *Mathematics of Neural Networks*, Oxford University, 1995.
10. Linear Algebra for Signal Processing, (edited with A. Bojanczyk), Springer-Verlag, New York, 1995.
11. Uncertainty in intelligent and transportable agent systems, preprint, 1996. (with K. Moizumi)
12. Mobile agents for mobile computing, (Robert Gray, David Kotz, Saurab Nog, Daniela Rus and George Cybenko), to appear in *Proceedings of 1997 Parallel and Distributed Architectures Symposium*, Aizu, Japan, March 1997.
13. Agent Tcl, In William Cockayne and Michael Zypa, editors, *Itinerant Agents: Explanations and Examples with CDROM*, Manning Publishing, 1996. To appear. (with R. Gray, D. Kotz and D. Rus)
14. 3D Base: Solid Object Recognition, to appear in *Journal of Smart Engineering Systems*, 1996. (with A. Bhasin and K. Cohen)
15. Neural Networks in Computational Science and Engineering, in *IEEE Computational Science and Engineering*, Spring 1996.
16. Robert S. Gray. Agent Tcl: A transportable agent system. In James Mayfield and Tim Finin, editors, *Proceedings of the CIKM Workshop on Intelligent Information Agents*, Fourth International Conference on Information and Knowledge Management (CIKM 95), Baltimore, Maryland, December 1995.
17. Robert S. Gray. Agent Tcl: A flexible and secure mobile-agent system. In Mark Diekhans and Mark Roseman, editors, *Proceedings of the Fourth Annual Tcl/Tk Workshop (TCL 96)*, Monterey, California, July 1996.
18. The Shannon Machine, In *Proceedings of the 9th Yale University Workshop on Adaptive Control and Learning* June, 1996. (with R. Gray, A. Bhasin, B. Brewington)

Personnel Supported by this Project

- Faculty: Professor George Cybenko

- **Graduate Students:**

1. Aditya Bhasin (MS completed 1996 present employer Booz-Allen)
2. Yunxin Wu (MS completed 1995 present employer Microsoft)
3. Robert Gray (Ph.D. expected 1997 summer employer AT&T Bell Labs)
4. Kurt Cohen (MS completed 1996 present employer McKinsey)
5. Katsuhiro Moizumi (Ph.D. expected 1998 summer employer Furukawa Electric)

- **Visitors:**

1. Professor Lennart Ljeung, Linkoping, Sweden visited and consulted about reinforcement learning at Wright Labs.
2. Professor Y.P. Wu, Beijing Institute of Technology, China visited to work on reinforcement learning methods.
3. Professor F. Preperata, Brown University visited to work on distributed information systems.